



Partecipazione e Conflitto
 * *The Open Journal of Sociopolitical Studies*
<http://siba-ese.unisalento.it/index.php/paco>
 ISSN: 1972-7623 (print version)
 ISSN: 2035-6609 (electronic version)
PACO, Issue 7(2) 2014: 258-277
 DOI: 10.1285/i20356609v7i2p258

Published in July 15, 2014

Work licensed under a Creative Commons Attribution-Non commercial-Share alike 3.0 Italian License

RESEARCH ARTICLE

THE DOWNS AND UPS OF THE CONSUMER PRICE INDEX IN ARGENTINA

From National Statistics to Big Data

Celia Lury

Centre for Interdisciplinary Methodologies, University of Warwick

Ana Gross

Centre for Interdisciplinary Methodologies, University of Warwick

ABSTRACT: On the 5th of February 2007, the Institute of National Statistics and Census in Argentina (INDEC) released a press statement, giving a percentage figure for that month's Consumer Price Index (CPI-GBA). Since the announcement, this number and its subsequent variations have been at the centre of a national and international political, legal and technical controversy. The legitimacy of the numerical value of the percentage has been called into question by a range of actors and has been challenged by the emergence of multiple alternative indicators of inflation. We explore this methodological controversy through the lens of stactivism. We do not describe the controversy in its entirety, but, rather, enter the controversy to develop a comparison of the procedures informing the production of the CPI as a national statistic with those informing its production as a big data number. In both cases, we explore the way in which price is produced as an indicator. In doing so we draw attention to the significance of calculative infrastructures as ubiquitous, multi-layered processes of connectivity, that have the capacity to make surfaces, to draw lines and boundaries, and to enable particular economic and political activities to unfold in multiple and specific ways. We argue that the capacity to connect, to attach and detach, that is immanent to such infrastructures configures price as an indicator in particular ways, and in doing so help make what we call state space, a term which we use to draw attention to how specific configurations of connectivity in the calculative infrastructure enacts a space of possibility for stactivism.

KEYWORDS: big data, consumer price index, controversy, experiment, national statistic, price

CORRESPONDING AUTHORS: Email: C.Lury@warwick.ac.uk; A.Gross@warwick.ac.uk

1. Introduction

On the 5th of February 2007, the Institute of National Statistics and Census in Argentina (INDEC) released a press statement:

The General Level of the Consumer Price Index for Capital Federal and the corresponding districts of Greater Buenos Aires registered in January a variation of 1.1% in relation to the previous month, and of 9.7% with respect to the same month in the previous year. (Pedido de Indagatoria 2007, 99).

Since then, this regional Consumer Price Index for Capital Federal and Greater Buenos Aires (CPI-GBA) and its subsequent variations have been at the centre of a national and international political, legal, and methodological controversy (Callon 1981; Latour 1987;

<http://www.mappingcontroversies.net>, last accessed on 16th April 2014). The legitimacy of the numerical value of the percentage has been called into question by a range of actors and has been challenged by a variety of both public and private alternative measures. We do not describe the controversy in its entirety here¹ but rather focus on two specific moments which are particularly salient to the exploration of *statactivism*, a term which, following Bruno, Didier and Vitale (2014), we use to refer to activities in which the methodological assumptions of statistics are opened up to political contestation. Our aim is to develop a comparison of the potential for statactivism in relation to the procedures informing the production of the CPI as a national statistic with those informing its production as a big data number.

Our approach to statactivism situates it within an understanding of state or problem spaces, a term which we adapt from Deleuze (2004) and DeLanda (2005)² to refer to the system or sphere of governance as a space of possible states, a problem space that supports diverse to-

¹ We use a range of empirical materials to inform the analysis we make here. Interviews were conducted with six individuals who were at the centre of the controversy in January 2007, including some of the statisticians dismissed from the INDEC. One interview was conducted with an external consultant to the government who has advised on methodological changes to be implemented as part of the intervention. The names of these interviewees are not included here to safeguard confidentiality. Additional materials include the following documents: Legajo de Causa Fiscalía de Investigaciones Administrativas 2007; Pedido de Indagatoria por parte de la Fiscalía de Investigaciones Administrativas 2007; Informe Técnico de la Universidad de Buenos Aires (UBA) Con Relación a la Situación del INDEC 2010; Manual Metodológico Número 13 Índice de Precios al Consumidor Gran Buenos Aires Base 1999=100; Manual de Índice de Precios al Consumidor – Teoría y Práctica, OIT; World Economic Outlook FMI (April 2010); and books by Jueguen F. and Bullrich L. (2010), *INDEC – Una Destrucción con el Sello de Los Kirchner*, Buenos Aires: Editorial Edhasa, and Noriega G. (2010), *INDEC – Historia Intima de una Estafa*, Buenos Aires: Editorial Sudamericana. The pictures we use form part of a photographic archive of INDEC's conflict produced by Claudio Rafael Castro.

² These writers themselves locate the term in relation to the history of thermodynamics, in which, in general terms, it refers to spaces of energetic possibilities or phase spaces. More specifically, in the theory of discrete dynamical systems, a state space is the whole set of values that processes can possibly take.

pologies of possibility and power³. The advantages of adopting the term for us is that it provides a conceptual framing for the empirical investigation of the potential for stactivism, emancipatory or otherwise, without requiring us to presume, in advance, who the actors involved in the controversy might be, or what interests or goals might be at play. This approach has much in common with that outlined by Stephen Collier in his discussion of Foucault's later work on biopolitics in terms of topologies of power (2009), with its emphasis on problem spaces and "the patterns of correlations" in which heterogeneous elements – techniques, material forms, institutional structures and technologies of power – are configured, as well as the "redeployments" and "recombinations" through which these patterns are transformed' (Collier 2009: 80). As Collier notes, the concern of such a mode of analysis is not the description of a multiplicity of contingent forms but rather the identification of the broad configurational principles through which new formations of government are assembled without recourse to the assumption of any inner necessity or coherence.

To make an empirical investigation within this mode of analysis, we situate our study of CPI in relation to a calculative infrastructure that configures the emergence of price as an indicator of a particular kind, specifically as an indicator that measures the variation of price. In doing so, we understand the calculative infrastructure to be organised in terms of multi-layered processes of ordering and connectivity that have the capacity to make surfaces or planes for action, to draw lines and boundaries, and to enable particular economic, political and semiotic effects to unfold in multiple and specific ways (Bowker and Star 1999; Thrift 2005; Day, Lury and Wakeford 2014). We argue that the capacity to connect and disconnect that is created in such an infrastructure is often what is at issue in state space in general and in this case in particular, in that its configuration informs the possible values in relation to which the emergence of price as a particular kind of indicator can acquire legitimacy. To empirically describe such a configuration in its entirety, however, would be a huge task, and so we focus here on two specific moments that allow us to develop a comparison of the procedures informing the production of the CPI as a national statistic with those informing its production as a big data number. This, in turn allows us to consider these two moments in terms of stactivism.

The first moment we address is described in Argentina as 'the intervention'. In its most literal sense, the intervention refers to the moment in the controversy in which long-standing statisticians were dismissed from the INDEC and substituted by others, more closely aligned with the government. However, the intervention may also be understood as a moment in which the legitimacy of the terms of the referentiality of price as an indicator was brought into question. As this referentiality was questioned so was the indicator's capacity to represent brought into doubt. In short, the intervention was a moment in which the controversy was revealed as having explicitly representational as well as political dimensions. To explore the significance of this moment in terms of state space we describe (some of) the relations, the degrees of connectivity and specification of entities, the processes of attachment, detachment and re-attachment of

³ For an empirical articulation of state spaces via controversial instances see for example Fraser (2009).

data within the calculative infrastructure that was deemed necessary to allow price to emerge as an indicator that had legitimacy as a national statistic. We then turn to a second moment within the controversy: the proliferation of alternative measures of inflation, and conduct the same kind of analysis to look at how one of these alternative measures, specifically that produced by the Billion Prices Project (BPP), re-configures price as an indicator in the online environment of big data. In each case, we consider how the infrastructural configuration of the CPI enacts a space of possibility for stactivism.

2. Controversial Referentiality: The Values of Price

To start though, let us briefly introduce the CPI: a widely accepted definition suggests that it is an indicator of the average measure of change in the prices of a set basket of consumer goods and services bought for the purposes of consumption by a representative average of households of a specific geographical territory. This territory is sometimes, but not always, the territory of a nation state: in the case we discuss here, for example, the CPI relates only to a particular region of Argentina – the Capital Federal and Greater Buenos Aires region. As such the geographical referent of this CPI is not the country as a whole. Nevertheless, the territory to which it refers is established as an administrative region of the Argentine nation, and it is the Argentine state that has responsibility, and a large administrative apparatus, for the production of this indicator on an ongoing (monthly) basis.

An important point to make here is that the CPI is an example of a composite number (Guyer 2010); that is, the CPI signifies – in a single number - an ensemble of heterogeneous things that exist in reality in a differentiated manner (Desrosières 1998). Moreover, it is a measure that transforms a complex phenomenon into a single value that varies in time (Bruno, Didier and Vitale 2014). Indeed, before the controversy, CPI-GBA was produced so as to be a ‘pure price index’ (Manual de Metodología Número 13, 2001): ‘pure’ being used here to suggest that only *changes in prices* between different reference periods are what is to be expressed in the index. This process of purification requires the isolation of price factors in a fixed (base) period of time, a comparison with the same factors in a second period, and then an estimation of the change that occurs if everything else – such as consumers and their spending patterns, the market and the goods and services it offers, the quality of the goods and services on offer, the condition in which products are encountered and so on – is taken out of the equation. Price factors are only ‘those which identify the amount paid for one unit of a specified kind and quality of good or service obtained from a specified source by a specific population group’ (Hurwitz 1962, 815). What is involved in this process of purification then is that any variation in price observed is not contaminated by other kinds and modes of variation. Indeed, it is the task of the CPI methodology to render such other modes of variation demonstrably *external* to any observable variation in price. We discuss some of how this is done below but it is important to recognize that in addition to being a product of this process of purification, a CPI is also typically designed to be put in relation to another complex phenomenon: money – or rather, it is designed to be put in rela-

tion to a specific capacity of money, that is, the 'purchasing power'⁴ of a currency. Indeed, it is by putting the observed variation in price of products that comprise the everyday activity of consumption by a population of households inhabiting a specific geographical territory in relation to international rates of currency exchange that the composite figure of the CPI brings the economy, state and society into relation with each other.

Having made this preliminary description, let us now turn to our first moment in the controversy. Throughout 2007 in Argentina, a series of public denunciations⁵ of the numbers presented as CPI-GBA led to the setting up of an independent legal investigation to be carried out by the Fiscalía de Investigaciones Administrativas (FIA)⁶ to determine whether the CPI-GBA had been manipulated in an inappropriate way, and specifically, to ascertain whether CPI-GBA made adequate, representative and reliable reference to the fluctuation of prices in the region (Figure 1). The public denunciations were initially mostly concerned to object to the government's intervention into the activities of the INDEC but they also questioned the accuracy and epistemological validity of the CPI-GBA produced by the now intervened INDEC (Figure 2). A variety of actors alleged that the normal procedures for the production of CPI-GBA's historical series had been altered without introducing a corresponding clear-cut change in the methodological guidelines for the collection, aggregation and weighting of prices⁷. Demonstrators also declared that primary data had been deleted and replaced with secondary data, external to the measuring procedures and sampling techniques traditionally deployed in the production of the CPI-GBA at the time of the controversy⁸. In addition, it was further questioned whether the data used for

⁴ The Office of National Statistics in the United Kingdom notes that 'the amount of money needed to purchase a fixed basket is also known as the internal purchasing power of the currency, which can be expressed in two ways. Firstly, it is the amount of money needed in period y to purchase the same basket of goods and services that one pound could purchase in an earlier period x . Conversely, it is the amount of money needed in an earlier period x which could buy the same basket of goods and services that one pound purchases in period y .' (ONS Consumer Price Index Technical Manual 2012, 2).

⁵ The initial denunciations were made by statisticians working at the INDEC at the time of the intervention, denunciations that were covered in all the main newspapers in the country (La Nación, Clarín, Perfil and Página 12). It is worth noting that the Asociación de Trabajadores del Estado (ATE), a branch of the Central de Trabajadores de Argentina (CTA) was particularly supportive of such denunciations. The FIA received two requests as part of the investigation: the first one was filed by the Asociación Civil Sin Fines de Lucro 'Asamblea por los Derechos Sociales' led in 2006 by Dr. Aníbal I. Faccendini. The second one was filed on the 14th February 2007 by a number of national senators.

⁶ The FIA is an independent organ of government dedicated to the investigation of corruption and administrative irregularities undertaken by agents within the national administration of government. The FIA mainly undertakes preliminary investigations, and facilitates the filing of legal cases.

⁷ For example, it was alleged by our interviewees that the database that was used to produce the Index was altered so as to allow – inappropriately – the limits on price change agreed by the government with some companies in relation to some categories of goods – to inform the calculation of the CPI.

⁸ Our analysis of legal materials and the interviews we conducted suggest that the secondary data collected and produced by the Secretary of Commerce and not that collected by INDEC was included as part of the production of CPI-GBA for January 2007. See Berumen and Becker (2011) for further discussion.

the elaboration of CPI-GBA was produced in accordance with the terms of the legal framework designed to preserve national statistics and their peculiar relation to the real.

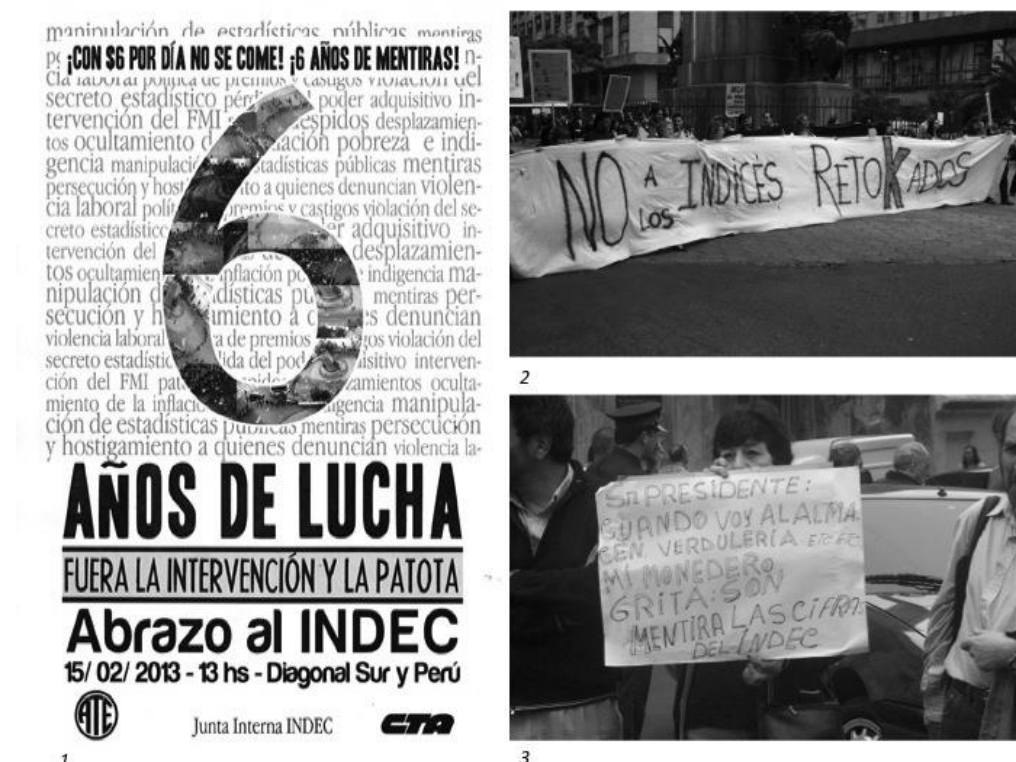
Figure 1 – Documents taken out of the INDEC for Legal Investigation, 25th July 2007



Source: *Photographic registry of INDEC's conflict by Claudio Rafael Castro*

Here we focus only on this last set of concerns, specifically the queries relating to Law Number 17.622 of 1968, article 10 and 13 concerning the need for statistical secrets. The legal documents which emerged from FIA's enquiry and the interviews one of us conducted with statisticians and technicians dismissed from INDEC following the intervention suggest a detailed disaggregation of the prices and objects of measures taken into account to construct CPI-GBA was demanded of INDEC by members of government. This included a request to the professional statisticians employed at INDEC to provide details of the specific retail outlets and brands of products being taken into account in the production of the CPI-GBA. This request was refused.

Figure 2 – Demonstrations



Some examples of the public demonstrations that followed the Intervention in 2006: 1 – A poster produced by the Union, ATE/CTA, inviting members of the public to participate in an Embrace of the INDEC building. The poster reads: ‘It Is Not Possible to Eat with 6 Pesos a Day! 6 Years of Lies! 6 Years Fighting! Out the Intervention and the Gang.’ 2 – A photograph of ATE’s first protest on April 10th, 2007. The banner reads, ‘No to Retouched Indexes’. The capital K in ‘retoKados’ (retouched) refers to the first letter of the then President’s surname ‘Kirchner’. 3 – A protester at INDEC: ‘Mr. President, when I go shopping my wallet screams: INDEC’s figures are a lie’.

Source: Photographic registry of INDEC’s conflict by Claudio Rafael Castro

Before the intervention, the CPI-GBA was designed to select and specify products according to the following hierarchical classification: 9 Chapters (e.g. Food and Drink), 26 Divisions (e.g. Food), 65 Groups (e.g. Fruits), 123 sub-groups (e.g. Fresh Fruit), 182 Products (e.g. Citrus Fruits) and 818 Varieties (e.g. Oranges) (Manual de Metodología Número 13 2001). The selection of goods itself is decided on the basis of the INDEC’s identification of which goods might comprise those purchased by a representative average household. From within this selection, and within the terms of an internationally shared standard CPI methodology, goods are further specified:

they are described through the listing of particular attributes such that they – or their substitutes – can be reliably identified in reality. This classification involves specifying not only characteristics of the goods as such, but also aspects of the way they are produced, their packaging, their distribution pathway or country of origin. It is thus a complicated process of specification, resulting in an indexical action of ‘pointing to’, sometimes culminating in the naming of a specific branded product held to exemplify these attributes of the item whose price was to be observed (Figure 3).

On the 11th of July 2006, Guillermo Moreno, then Secretary of Commerce, made the first of a number of formally documented requests for information to the INDEC, asking for the release of data which would reveal not just the Variety, but the further specification of the products of which prices were collected to produce CPI-GBA. He also requested the details or geographical coordinates of the retail outlets at which the prices of such specified products were being observed. Government officials justified the request by saying that the release of the information was necessary so that the government could face ‘the serious problem that inflation represented for the country, manage the inflation expectations of the population and the impact these might have over the future economic development of the nation’ (Guillermo Moreno – Secretary of Commerce, *Pedido de Indagatoria* 2007, 1; our translation)⁹. Despite repeated requests, the INDEC statisticians refused to disclose information to the government about products to this degree of specification.

Until the so-called intervention, the INDEC had routinely made public changes in prices to the level of a product’s Variety¹⁰, that is, it made public average price variation to the level of a relatively unspecific generalisation: for example, Oranges. When questioned at the legal enquiry, one of the INDEC statisticians responded to the request for more specific information by saying, ‘if we had released information of this kind [that is, at the level of specific products in specific places] to Moreno, we would have handed him the elements to regulate and control the prices of the specified products and that would have had an effect on the Index’ (*Pedido de Indagatoria*, 29-30; our translation). The reason given by the statisticians for the refusal to give the information was thus the desire to prevent government intervention in the production of the CPI at the level of the manipulation of the specific prices of specific products. To justify the refusal,

⁹ The request for this information by the government was widely interpreted to be part of an attempt to manipulate the CPI-GBA. While some commentators also argued that such requests were motivated by the desire to correct the perceived plutocratic bias reproduced in the Index, others suggest that the government’s intention was actually to decrease the value of debt bonds linked to the CER (Reference Stabilisation Coefficient), which involves the use of the CPI-GBA.

¹⁰ In contrast, some other countries provide figures relating to products more closely specified. So, for example, the UK Office of National Statistics (ONS) recently released what is defined as ‘CPI’s micro-data’ following a Freedom of Information Act request that was filed in 2011 demanding the publication of individual item indices. Since then, ONS has extended the network of attributes that make visible each of CPI’s specified products as part of its Open Data agenda. However, because of other regulatory frameworks still in place, products are often still unidentifiable in their actualized forms, in large part because brand names are excluded from public data releases in the UK.

the INDEC's statisticians referred to Law Number 17.622 which sets out the terms of the regulatory framework by which it is required both that some of the specific characteristics of data, and that the combinatory possibilities that may be established between different data units in the production of knowledge for public disclosure be kept secret¹¹.

As such, the refusal might be seen to have merited widespread support (which indeed it did secure) insofar as the statisticians can be seen to be protecting the interests of the public as a whole by seeking to prevent the manipulation of the data for political motives. But we want to stop here for a moment, and consider what else – in addition to the desire to prevent the 'drawing' of numbers¹² – is involved in this refusal and ask what else this refusal might reveal about the potential for stactivism within the configuration of the calculative infrastructure that produces the CPI as a national statistic. To do this, we need to recognise the historical and technological contingency of price as an indicator, and acknowledge the variety of ways in which it emerges empirically as part of particular orderings of the flow of information (Pardo-Guerra 2010; Muniesa 2007; Beunza, Hardie and Mackenzie 2006). Muniesa, for example, argues that as entities prices are eclectic; they may be 'agreement[s] to close a transaction between two parties, a tag on a grocery shelf, an expert's indicative estimate of what a fair value of something should be or a synthetic input for a mathematical formula. Prices can be found or they can be proposed. They can be taken for granted or observed with expectation. They can be regarded as fair or unfair, as accurate or inaccurate, as good or bad, but also random, meaningless events' (2007, 378). Guyer (2006) forcefully insists that price is itself a composite, emergent in processes of composition, including, for example, calculations of risk. Interestingly she also suggests that this compositeness is becoming more widely understood by the general public, as a variety of processes of composition become more visible¹³. It is no surprise then that the operations we describe above comprise a complex series of connections, transformations and in-

¹¹ As a result of the statisticians' refusal to provide these details of data collection to the government, a Committee was formed to determine what data should legally be disclosed in response to requests by the government. The Committee's conclusion was in accordance with the Act, namely that data could be disclosed only insofar as: a) it does not enable the identification of price informants, be these physical location(s) or person(s); and b) does not allow for the identification of a brand or reveal the actual (real) object or substance of measure. Identifying yet other actors whose political interests in the CPI were revealed to be at issue by the controversy, it was held that the removal of the legal barrier could 'make vulnerable the legitimate interests of producers and of informants' (Pedido de Indagatoria 2007, 37; our translation).

¹² In Argentinean parlance, a false number is commonly described as a number that has been drawn. Drawing a number can entail a multiplicity of actions: guesstimating, manipulating, fiddling, tinkering, distorting, juggling, inventing. It refers to practices by which numbers emerge that are deemed to be non-calculative in nature. The expression 'drawn numbers' thus points to a representational fissure or crack between a number and that to which they should objectively refer – in this case, changes in the pure price of consumer goods and services.

¹³ To take an example in the UK, there is a currently a dispute between energy companies, the government and environmental groups about how much of the ever-increasing price of domestic energy is due to a government charge to support alternative energy measures added to all bills.

terventions in the calculative infrastructure and produce price as a very specific kind of an indicator, necessarily requiring, as part of its production, that some relations to the world be preserved and others lost (Latour 1999)¹⁴.

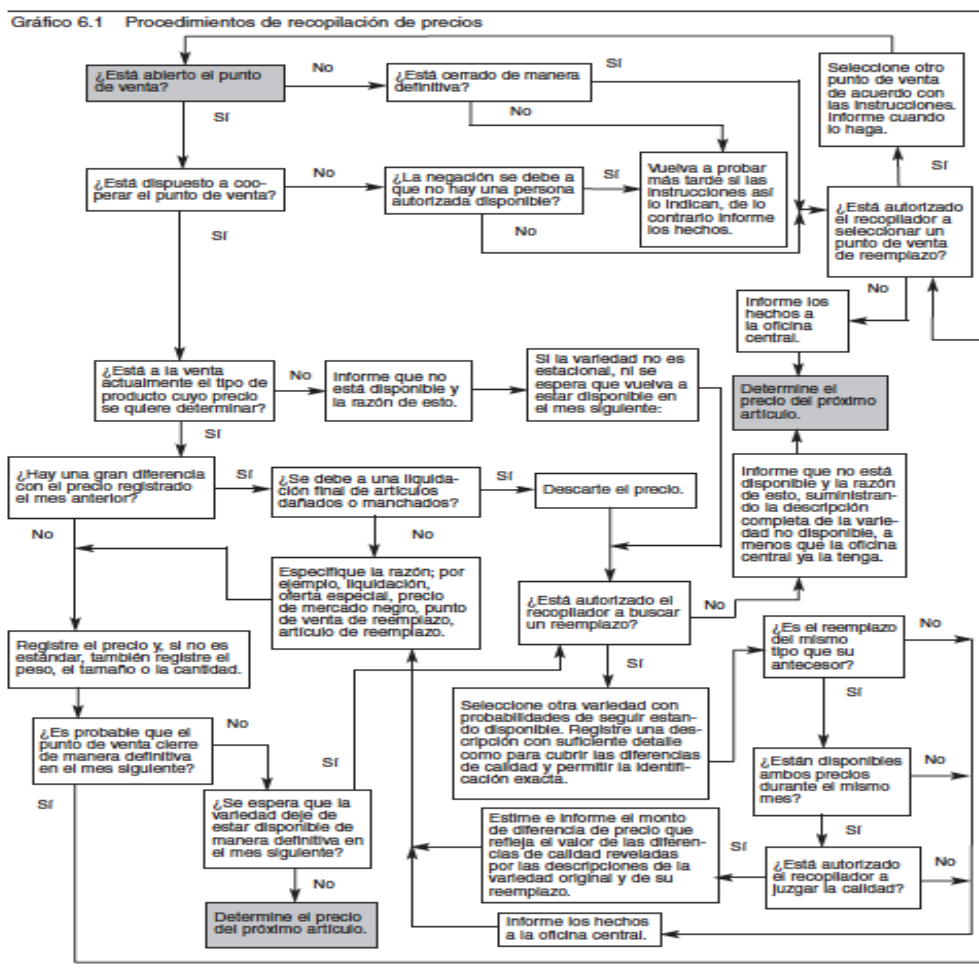
Indeed our suggestion is that there is a complex configuration of the calculative infrastructure at work in each and every production of a CPI, involving not only the establishment of links, the mapping of territory, and the specification of goods but also the creation and maintenance of a whole range of other methodological connections and disconnections. And the refusal of the statisticians to release the information requested of them by the government described above certainly demonstrates how the production of CPI-GBA as a legitimate measure was seen by the statisticians themselves to do so. Their action demonstrated their belief that a 'sphere of obscurity and a structure of withholding' is required for the legitimacy of the CPI-GBA: their actions were directed to the protection of a (state) space of secrecy from which 'the data could first be developed and from which the figures could emerge as part of its process of entering the public sphere' (Didier 2005, 639)¹⁵. But it also cut across another plane of activism, that involving challenges to their own expertise in selecting and specifying goods in reality in terms of their representativeness.

Indeed in the context of this moment in the controversy the refusal by the statisticians to release this information points to the importance of the legal restriction on the flow of information not simply as a barrier to manipulation but as an interface within state space, able to be operated by the statisticians, so as to enable a particular directionality in the flow of information, resulting in a kink or fold in the relation between levels of observation within the calculative infrastructure. This directionality enables information to flow so that actual products may be identified by the statisticians so that changes in their price can be directly observed as part of their calculation of the CPI, while simultaneously allowing the prices of those products to be detached or disconnected from their socio-technical, political environments for all actors other than the statisticians. In this way, the operation of the legal restriction as an interface by the statisticians allows for price to be made available for scientific or statistical observation, to have a real context, but does not allow others to demand that alternative observations of that real context be allowed to enter the calculations required for the production of the CPI.

Figure 3 – Specification of Products for CPI

¹⁴ Latour argues that 'it seems that reference is not simply the act of pointing or a way of keeping, on the outside, some material guarantee or the truth of a statement: rather it is a way of keeping something constant through a series of transformations.' (1999, 58)

¹⁵ That is, it sought to protect the details of their first order observation from any second order observation other than their own (Esposito 2013; Stark 2013).



Source: OIT Graph 6.1: Procedures for the Collection of Prices (OIT Manual 2006, 103)

By refusing the request for information the statisticians thus prevented ‘experiments in contexting’ (Asdal and Moser 2012); that is, they both prevented the government or other groups from manipulating prices directly and prevented the government or other actors from challenging the particular relation established between representation and representativeness they held together as a pure index – on behalf of the public – until the intervention.¹⁶ In doing so, they

¹⁶ Indeed, some have suggested that the reason the government requested information from the statisticians was to correct the plutocratic bias¹⁶ that CPI-GBA was deemed to deploy as part of its calculus; that

both opened up and closed down the possibilities of statactivism that are distinctive to the production of the CPI as a national statistic.¹⁷ The variation in the price of *generic* Oranges differs from the variation in the price of Organic Navel Oranges sold in gentrified neighborhoods of Capital Federal, to the price of residual Osage Oranges sold in the popular markets of some Greater Buenos Aires districts. But if the information about which *specific* Oranges are observed for changes in their price is not revealed then not only is the government or other agencies prevented from manipulating prices but also the relation between the referentiality of the index and its representativeness in relation to specific localities or specific income groups is kept hidden, removed from the political realm, and reserved as a technical decision.

To conclude this section then: how and what information flows, and in which directions, is revealed in this moment in the controversy to be what organizes the possible dimensions of statactivism. In this case, the withholding of information required by law, upheld by the statisticians, sustained the processes by which the measurement of price is protected from manipulation at the level of the observation of specific products *and* allows statisticians to render price as a pure index, a referring simultaneously to the real and to nothing but itself (in a previous time period). Our argument has been that this action reveals the forms of statactivism peculiar to the production of the CPI as national statistic: if the operation of the barrier as interface had been lifted, the configuration of price as not simply a measure but as a value would have potentially been at issue. It would have become possible both to manipulate the prices observed and to contest whether the variation in price shown by the CPI related to a variation in time for goods and services that was representative of the consumption patterns of the whole population of reference¹⁸ or of some groups more than others.

3. Billion Prices, Alternative Measures

Following the publication of the findings of the investigation carried out by the FIA, the controversy over the validity of the CPI-GBA continued to grow and many other national indicators were challenged, including, for example, measures of poverty and extreme poverty, real earnings, real exchange rate, GDP, and so on (Lindenboim 2013; Becker 2013; Berumen and Becker 2011). Nevertheless some provinces continued to estimate and release to the public their own

is, the government request for this information was seen by some to be motivated by a desire to show that the specification of products being selected by the INDEC for the calculation of the Index was not representative of the consumption patterns of the population as a whole. See for example www.indec.com.ar, last accessed on 16th April 2014.

¹⁷ In suggesting that such forms of statactivism are a distinctive possibility in relation to the production of national statistics we are not, of course, saying that they are the only forms possible.

¹⁸ For IPC-GBA the population of reference is the one estimated in the Encuesta Nacional de Gastos de los Hogares (ENGH).

regional CPIs¹⁹, with some of such numbers coming to be regarded as the 'real' measures of inflation in the country. At the same time, rival measures to the CPI-GBA were also developed by academic economists and representatives of various private institutions²⁰, who developed independent, non-governmental projects to observe the prices of a basket of goods and services in order to calculate inflation rates. These typically produced a figure of variation in prices significantly higher than the official estimates produced by the INDEC for Capital Federal and Greater Buenos Aires.

The response of the government to these alternative measures was to ban them – by fining the organizations producing them²¹ – on the grounds that they were misinforming the consumer. Indeed, the government suggested that the only entity with the methodological capacity to assess the CPI rigorously at the geographical level of either the Capital Federal and Greater Buenos Aires or the nation is the INDEC; they asserted that such a complex, large-scale operation is economically and logistically impossible for private organisations. In response to the ban by the government, the private consultancies came together in June 2011 under the banner of *IPC Congreso*, an initiative by which members of the opposition (including some of those belonging to the Commission for Freedom of Expression in the Cámara de Diputados) made the figures produced by these private consultancies public in the Congress in order to avoid being fined. However in October 2013, the Supreme Court ruled that the government fines were illegal and private consultancies began to publish their inflation figures independently again. More recently still (January 2014), the government released a newly designed national CPI (IPCNU) that (implicitly at least) addresses the technical and political flaws identified at the time of INDEC's intervention. In its first release the new indicator, which has been welcomed by a range of actors, including the media, economic experts and international bodies, has resulted in an inflation figure that is widely deemed to more accurately express the real.

In this section of the paper, however, we want to look at one particular spin-off measure linked to an earlier moment in the controversy that also provides an alternative inflation figure: the Billion Prices Project (BPP; <http://bpp.mit.edu/>, last accessed on 16th April 2014). This pro-

¹⁹ Dirección Provincial de Estadística y Censos, Provincia de San Luis (DPEyC – San Luis); Dirección Provincial de Estadísticas e Investigaciones Económicas, Provincia de Mendoza (DEIE – Mendoza); Dirección General de Estadística y Censos, Provincia de Córdoba (DGEC – Córdoba); Instituto Provincial de Estadística y Censos Provincia de Santa Fé (IPEC – Santa Fé); Dirección General de Estadística y Censos, Ciudad Autónoma de Buenos Aires (DGEC – Capital Federal).

²⁰ Buenos Aires City, Gabriel Rubinstein & Asociados, Estudio Bein & Asociados, EconViews, M&S Consultores, Ecolatina, FIEL, Abeceb.com, Asociación de Dirigentes de Empresas (ADE), Finosport and Orlando Ferreres y Asociados among others.

²¹ In 2011 the Secretary of Commerce Guillermo Moreno fined private economic consultancies on the basis of an alleged violation of article 9 of law number 22.802. This Law stipulates that any public presentation, publicity or propaganda containing inexactitudes leading to consumer deception or confusion may be sanctioned. It was considered by the government that the publication of alternative inflation indexes that lacked scientific rigour could induce confusion in members of the public, leading them to make ill-informed decisions and that such publications should therefore be banned on the basis of this Law.

ject provides a challenge not simply to the particular figure that emerges as the percentage variation in price put forward as Argentina's CPI but also to the production of the CPI as a national statistic at all, since it involved the use of on-line or big data²² rather than the primary data of statisticians. The project was initially created by Alberto Cavallo and Roberto Rigobon, two economists, now based at the MIT Sloan School of Management in the USA. It currently monitors the daily price fluctuations of approximately 5 million items sold by approximately 300 online retailers in more than 70 countries to produce, among other things, a daily price index and inflation rate for a number of countries, including Argentina. The BPP originated in 2007 as part of Cavallo's PhD thesis²³ at Harvard University which compared the online price variations for Argentina, Chile, Brazil and Colombia. Cavallo later went on to produce an alternative online inflation index for Argentina called 'True Inflation' which was accessible via a webpage²⁴ and updated price variations for Argentina on a daily basis. In 2008, data collection was expanded to 50 other countries, an exercise which officially became the BPP. What started as an academic exercise, however, was later subsumed in a company that describes itself as 'the leading source of daily inflation statistics around the world', trading under the registered trademark of PriceStats²⁵. Indeed, the BPP website now refers those seeking 'more high-frequency inflation data across countries and sectors' to PriceStats, 'the company that collects the online data we use in our research initiatives and experimental indexes'.

As indicated above, what is interesting for us about the methodology of the BPP is that it made extensive use of online price data, taken from the largest supermarket chains, which it collected using scraping software. As a method, scraping is substantially different to the procedures used to specify and find products off-line in order to observe their prices.²⁶ While the production of CPIs using established statistical methodology requires the repeated, and constantly varying involvement of human agents (price collectors) and devices (questionnaires, flowcharts, instructions, etc.), scraping performs similar but much more automated set of operations over a relatively frictionless surface in the calculative infrastructure, that organized by HyperText Markup Language (although it will shortly switch to using Web code). Its scraping methodology is described by BPP as a three-step process. First, there is the downloading of se-

²² We do not want to provide a definition of big data ourselves, but one influential definition identifies three characteristics: volume (the amount of data generated), variety (the multiplying of, and ability to cross-reference, various types and sources of data), and velocity (the speed at which data is flowing) (Zikopoulos et al. 2012, 5-9). It is these interconnected transformations – the 'three v's', as Zikopoulos et al refer to them – that lead some commentators to suggest that 'a radically new kind of "knowledge infrastructure" is materializing' (Bollier 2010, 1) with far reaching implications for a range of spheres such as commerce, science, society, etc.

²³ See Cavallo A. (2010), *Scraped Data and Sticky Prices*, MIT Sloan.

²⁴ See <http://www.inflacionverdadera.com/>, last accessed 11th June 2013.

²⁵ See <http://www.pricestats.com/>, last accessed 29th October 2013.

²⁶ Indeed, it could be argued that the scraping methodology employed by BPP is only similar to offline methods for the creation of CPIs insofar as it involves a repurposing of off-line data (Marres and Weltevrede 2013) to inform its collection of the prices of commodities on- and –off-line.

lected web-pages where product and price information is given. Second, the underlying code of those web-pages is analysed by software to locate each piece of relevant information: 'this is done by using custom characters in the code that identify the start and end of each variable, according to the format of that particular page and supermarket' (Cavallo 2010, 7). Third, scraped information is then stored in a database containing 'one record per product per day' (Cavallo 2010, 7).

The nature of the data collected through the method of scraping thus differs substantially to that conventionally used in the production of CPI as a national statistic: first, the data is collected on a daily rather than monthly basis; second, the data collected is relevant for a number of countries, not just one, or not just one region within a country; third, the data comprises detailed information on the full array of a retailer's products²⁷, rather than being a selection specified in terms of measures of representativeness; fifth, there are no forced item substitutions – new products may simply be added; and sixth, the data are rendered comparable across countries, with prices recorded for the same categories of goods and time periods in different places. As a consequence, while the BPP can be used to confirm the belief that the official offline inflation figures in Argentina are inaccurate²⁸, it is far from being just a substitute methodology for the production of a CPI. Indeed, because the data collected is what could be described as naturally occurring – that is, it is not produced in response to directed (statistical) enquiry – but is, rather, 'born digital', its collection can, in principle, be more variously organised than can the data required for the production of national statistics. As such, it is a new configuration of power and possibility of state space within the calculative infrastructure and is, we suggest, transforming the possibilities of stactivism.

To describe this data as 'born digital' is not, of course, to say that its collection is without method or that its recording and collection is merely automatic. In the context of the controversy a whole series of contrasts may be drawn with the methodology that produces the data employed to produce the CPI as a national statistic. One important contrast is that while the BPP methodology allows specific prices to be mapped onto specific products in a one-to-one indexical relationship and products to be categorized, the categorization – for whatever purpose – comes after or *follows* the collection of data relating to prices and products from selected websites. As a consequence, although the BPP follows traditional CPI classifications and respects official weightings of prices for the purposes of producing a measure that is comparable to conventional CPIs, its methodology invites the experimental exploitation of the data through the use of multiple (and possibly changing) kinds of categorization, so allowing different kinds of

²⁷ This is important to the BPP since the ability to identify products displayed next to each other plays a key role in measuring price synchronization among close substitutes in economics.

²⁸ Cavallo notes, 'amongst their many potential uses, price indexes constructed with online data can be used to obtain alternative inflation estimates in countries where official estimates have lost their credibility. In particular, this paper uses online prices to evaluate the widespread claim that the Argentine government has been manipulating the official inflation index since 2007' (Cavallo 2012, 1).

comparison to be made across these multiple categories²⁹. In addition, the time periods over which price can be demonstrated to vary in relation to any and all of these categories can be intensively differentiated, allowing for both very short-term and long-term comparisons. Together, so we suggest, these differences in methodology produce a different surface, supporting the production of a variety of indicators of different kinds of relevance to different kinds of actors.

It is interesting in this respect to observe that the data collected by PriceStats by this scraping methodology is made available to others in commercial, academic and 'public' forms, defined by varying terms of access. Thus PriceStats website says that it: distributes (that is, presumably, it sells) its daily inflation statistics through an exclusive partnership with State Street Global Markets (<http://www.statestreetglobalmarkets.com>, last accessed on 16th April 2014), whose target clients include hedge funds, pension funds, and sovereign wealth funds; is in partnership with the BPP, described by PriceStats as 'an academic initiative that uses high frequency price information to conduct breakthrough economic research'; and collaborates with public institutions to improve decisions in public policy: for example, they create special indices that measure the price of specific goods across countries to 'anticipate the impact of commodity shocks on low income, vulnerable populations'. In short, the scraped data is used by PriceStats for very many different kinds of calculative operations, in relation to specific kinds of clients, resulting in, for example, the production of measures of real-time global inflation, price synchronization among brands and contiguous products, price stickiness, price markups, etc. In this regard it appears as if its data has the potential to be used not only to challenge the legitimacy of the CPI as a national statistic in particular countries or nation states but also to offer new opportunities for the commercial operation of the purchasing power of money. Its re-configuration of the calculative infrastructure may thus be seen to be enabling what might be called the financialization of inflation insofar as it provides the data for economic actors to put multiply categorized variation in prices in relation to the purchasing power of money, with the possibility of detaching the change in prices observed from any normative or exclusive relation to wages, pensions, savings, and the reproduction of a population in any particular country³⁰. Rather than being tied to categories produced in relation to certain kinds of existential indexicality specified in measures of representativeness, it invites *experiments in categorizing* within data sets in relation to changeable (typically macro-economic) categories. Such categories are no longer required to be created in relation to measures of representativeness for the population of specific societies, but can, rather, be created in terms of the opportunities they offer for political action or economic

²⁹ This is also increasingly true of open off-line data, although there are still limits on the level of specification to which such openness applies. In addition such data is also only available in relation to relatively limited time periods.

³⁰ The issue of referentiality does not of course entirely disappear; it can return in relation to the question of which websites are identified for scraping. And PriceStats themselves keep this information secret, not it would seem as a way of securing trust in the figures they produce (since it has the opposite effect), but so that others cannot duplicate their methodology.

advantage by specific actors. This is the new plane for stactivism afforded by the production of macro-economic indicators as big data numbers.

4. Conclusion

In this paper we have considered the relations between indicators, economics and politics in the context of the ongoing controversy about the CPI in Argentina, presenting an analysis that sought to demonstrate how two moments in this controversy show how national statistics and big data are implicated in the articulation of state space. Our analysis of a first moment in the controversy showed that the production of the CPI as a national statistic requires a specific (topographical and topological) spacing of state space that emerges through the ways in which, for example, a flow chart is mapped onto specifically qualified, temporally and geographically distributed, products, leading to the recording of data for use in calculations, the details of whose referents were folded into the calculative infrastructure, and protected behind a barrier of secrecy. What was revealed to be hidden (from manipulation and from contestation) was how the cost of living was mapped onto livelihood. Our analysis of a second moment identified the reconfiguration of state space involved in the production of CPI – and other macro-economic indicators – as big data numbers.

In both cases, state space is revealed to be dynamic, its continually changing configuration – in various processes of connectivity and dis-connectivity, of transparency and opacity, openness and closedness – providing the constantly shifting methodological and epistemological battle-grounds of stactivism. In the first, however, the values of the methodology are organized in terms of representativeness, operationalized through certain (protected) techniques of referentiality, while in the second, there are no a priori limits placed on the values; these may be established by clients, and they do not need to be linked to measures of representativeness in relation to populations in specific territorial spaces. In this second space, stactivism may be understood in terms of experimentation in categorizing rather than experimentation in contexting distinctive to the first. Of course, neither of such kinds of experiments – of contexting or categorizing – can operate independently of the other; it is simply that in the first moment, the organisation of the calculative infrastructure invites contestation in relation to the former, by making some actors more visible than others, while in the second, it is exercises in categorisation that appear to offer most opportunities to manoeuvre to those who have economic or political access to the data.

In this regard, we find it intriguing that, just as we came to the end of writing this paper (October 2013), there was an announcement of a new economic data-gathering entity: Premise³¹.

³¹ http://www.premise.com/introducing_premise.html, last accessed 29th October 2013. It is based in San Francisco and is backed by Google Ventures, Andreessen Horowitz and Harrison Metal. Its founder, David Soloff, used to be on the advisory board of Columbia's Institute for Data Sciences and was founding CEO of Metamarkets, a data analytics firm.

This new agency describes itself as a response to finding that '[o]ur experience at the grocery store had nothing in common with the inflation headlines we were reading in the paper. We wondered, what happens when you replicate and scale this divergence across ever-larger groups of people?' Its mission is 'to create faster, more accurate, more liquid data sets to deliver the critical transparency our electronic age demands'

(Premise, http://www.premise.com/introducing_premise.html, last accessed on 29th October 2013). Its methodology is described in one newspaper report as 'a blend of Google Street View and the CPI', or, alternatively, as 'an international photo-collage project'

(<http://www.theatlantic.com/business/archive/2013/10/the-price-network/280553/>, last accessed on 14 April 2014) and involves the collection of proprietary data from a number of what are defined as 'unique sources', including e-commerce websites, third-party transaction streams and their own mobile workforce which visits 'real-world' stores in Argentina, Brazil, China, India and the United States to 'ground truth directly' with the use of a mobile app that allows employees to photograph products with a time and date stamp. Perhaps this is a methodology that will enable experiments in both contexting and categorizing data. What will be interesting to see is whether and how it will contribute to the making visible of processes of re- and de-composing not simply the CPI but also price, and in doing so, what other dimensions of stactivism it may afford.

References

- Asdal K. and Moser I. (2012), "Experiments in Context and Contexting", *Science, Technology and Human Values*, Vol. 37 (4): 291-306, doi: 10.1177/0162243912449749
- Becker V. A. (2013), "The 2010 Census and the "Militant Statistic", *Radical Statistics* Issue 108: 56-57.
- Berumen E. and Becker V. A. (2011), "Recent Developments in Price and Related Statistics in Argentina", *Statistical Journal of the International Association of Official Statistics* 27: 7-11.
- Beunza D., Hardie I. and MacKenzie D. (2006), "A Price is a Social Thing: Towards a Material Sociology of Arbitrage", *Organization Studies* Vol. 27 (5): 721-745, doi: 10.1177/0170840606065923
- Bollier, D. (2010), *The Promises and Peril of Big Data*, Washington: The Aspen Institute.
- Bowker, G. C. and Star, S. L. (1999) *Sorting Things Out: Classification and its Consequences*, Cambridge (Massachusetts): MIT Press.
- Bruno I., Didier E., Vitale T. (2014), "Stactivism: forms of action between disclosure and affirmation", *Partecipazione e conflitto. The Open Journal of Sociopolitical Studies*, 7(2), doi: 10.1285/i20356609v7i1p
- Callon, M. (1981), "Struggles and Negotiations to Define What is Problematic and What is Not: The Socio-logic of Translation", in Knorr W.R., Krohn R. and Whitley R.P. (Eds.), *The Social Process of Scientific Investigation*, London: D. Reidel Publishing Company, pp. 197-221.
- Cavallo, Alberto (2010), *Scraped Data and Sticky Prices*, PhD Thesis, MIT Sloan.

- Cavallo, Alberto (2012), Online and Official Price Indexes: Measuring Argentina's Inflation, *Journal of Monetary Economics*, <http://dx.doi.org/10.1016/j.jmoneco.2012.10.002>
- Collier, S. (2009), "Topologies of Power: Foucault's Analysis of Political Government beyond 'Governmentality'", *Theory, Culture and Society*, Vol. 10 (6): 78-108, doi: 10.1177/0263276409347694
- Day, S., Lury, C. and Wakeford, N. (Eds.) (2014 forthcoming), "Numbering: ecologies and composition", *Distinktion* Special Issue, Number Ecologies: Numbers and Numbering Practices.
- DeLanda, M. (2005), *Intensive Science and Virtual Philosophy*, London: Continuum.
- Deleuze, G. (2004), *Difference and Repetition*, London: Continuum.
- Desrosières, A. (1998), *The Politics of Large Numbers: A History of Statistical Reasoning*, Cambridge (Massachusetts): Harvard University Press.
- Didier, Emmanuel (2005), Releasing Market Statistics, in Latour B. and Weibel P. (Eds.), *Making Things Public: Atmospheres of Democracy*, Cambridge (Massachusetts): MIT Press.
- Esposito, E. (2013) "Economic Circularities and Second-Order Observation: The Reality of Ratings", *Sociologica*, 2, Doi: 10.2383/74851
- Fraser, M. (2009), "Standards, Populations, and Difference", *Cultural Critique*, 71: 47-80, doi: 10.1353/cul.0.0032
- Guyer, J. (2006), "Composites, fictions and risk: towards an ethnography of price", in C. Hann and K. Hart (Eds.), *Market and Society: The Great Transformation Today*, Cambridge: Cambridge University Press, pp. 203-220.
- Guyer, J. (2010), "The Eruption of Tradition: On Ordinality and Calculation", *Anthropological Theory*, Vol. 10 (1): 1-9, doi: 10.1177/1463499610365378
- Hurwitz, A. (1962), 'Constants and Compromise in the Consumer Price Index', *Journal of the American Statistical Association*, Volume 57 (13): 813-825
- Latour, B. (1999), *Pandora's Hope: An Essay on the Reality of Science Studies*, Cambridge (Massachusetts): Harvard University Press.
- Lindemboim, J. (2013), Discourse, Legislation and Data in Argentina, *Radical Statistics* Issue 108: 54-56.
- MACOSPOL, Mapping Controversies of Science for Politics, <http://www.mappingcontroversies.net>, last accessed 29th October 2013.
- Marres, N. and Weltevrede, E. (2013), "Scraping the Social? Issues in Live Social Research", *Journal of Cultural Economy* Vol. 6 (3):313-335, doi: 10.1080/17530350.2013.772070
- Muniesa, F. (2007), "Market Technologies and the Pragmatics of Price", *Economy and Society* Vol. 36 (3): 377-395, doi: 10.1080/03085140701428340
- Pardo-Guerra, J.P. (2010), "Creating Flows of Interpersonal Bits: The Automation of the London Stock Exchange", c.1955-90, *Economy and Society* Vol. 39 (1): 84-109, doi: 10.1080/03085140903424584
- Stark, David (2013) 'Observing Finance as a Network of Observation', *Sociologica*, 2, doi: 10.2383/74854.
- Thrift, N (2005), *Knowing Capitalism*, London: Sage.

Zikopoulos, P., Eaton, C., DeRoos, D., Deutsch, T. and Lapis, G. (2012), *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*, eBook: IBM Corporation.

AUTHOR INFORMATION:

Celia Lury is Professor and Director and Centre for Interdisciplinary Methodologies, University of Warwick. She is currently studying the ways in which methods, metrics and models enact the social world. Recent publications include, *Inventive Methods: The Happening of the Social*, co-edited with Nina Wakeford, Routledge, 2012.

Ana Gross is a PhD Candidate at the Centre for Interdisciplinary Methodologies, University of Warwick and a visiting student at the Centre for the Study of Invention and Social Process (CSISP), Goldsmiths College, University of London. Her doctorate is currently funded by the Economic and Social research Council (ESRC).